

THESIS / THÈSE

MASTER EN SCIENCES DE GESTION

Quels sont les déterminants des retards au travail ?

Bendriss, Samira

Award date:
2018

Awarding institution:
Université de Namur

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



Travail de mémoire Master 60 Gestion HD

**Quels sont les déterminants
des retards au travail ?**

Bendriss Samira

Sous la direction de Monsieur Jean-Yves GNABO
Année académique 2017-2018

Table des Matières

1. Introduction	3
2. La problématique de la mobilité en Belgique	4
3. Revue de la littérature	8
4. Constitution de la base de données	10
3. 1 Élaboration d'un questionnaire.....	10
3. 2 Source des données	10
3. 3 Aménagement de la base de données	10
5. Présentation de la méthodologie et du modèle.....	11
5. 1 Etape 1–Analyse des Correspondances Multiples (ACM)	11
5. 2 Etape 2 – La Régression Logistique	11
6. Application de la méthode ACM	13
6. 1 Les hypothèses	13
6. 2 Choix des variables	13
6. 3 Choix du plan factoriel	13
6. 4 Caractéristiques des axes factoriels	15
6. 5 Interprétation	16
6. 6 Résultats	16
7. Application de la Régression logistique.....	17
7. 1 Les hypothèses	17
7. 2 Choix des facteurs à analyser dans la régression logistique	17
7. 3 Interprétation statistique	18
Statistiques descriptives	18
Coefficients d'ajustement	18
Test de l'hypothèse nulle H_0	18
Test de type II	19
Test de Hosmer-Lemeshow	19
Interprétation des paramètres du modèle	19
Tableau de classification pour l'échantillonnage d'apprentissage	20
La courbe ROC	21
7. 4 Interprétation économique.....	21
8. Conclusion – Limites de l'étude.....	23
Bibliographie.....	26
Annexes	30
Annexe 1 - Tableau des sources de données et des taux de réponse	30
Annexe 2 - Tableau synthétique des variables et de leurs instances dans la base de données.....	31
Annexe 3 - Graphique agrandi des modalités dans ACM	32
Annexe 4- Tableau des arrivées en retard pour les horaires mobiles, la direction et les cadres	33
Annexe 5- Détails des questions 10 et 12 extraites du questionnaire de départ	34

1. Introduction

L'intérêt manifesté pour le thème de cette étude est assez évident puisqu'il concerne la plupart des gens dans notre pays, les étudiants y compris. Arriver à temps aux est un défi auquel nombre d'entre nous sont effectivement confrontés au quotidien. On ne peut également qu'être interpellé par le peu de progrès apparent en matière de mobilité, et ce, en dépit des coûts très importants – économiques et humains – engendrés par le maintien de cette situation.

L'objectif de cette étude est de mettre en évidence les problèmes de mobilité propres aux différents moyens de transport et d'examiner si ces problèmes ont un impact direct ou indirect sur l'importance des retards. Il s'agit là d'un angle non abordé dans la littérature existante.

Après la présente introduction, les points suivants seront présentés : réflexion sur la problématique de la mobilité (chapitre 2) ; passage en revue de la littérature pour dresser l'état des connaissances sur les sujets et thème abordés dans ce mémoire (chapitre 3) ; présentation des données avec mention de leur source, détail de la procédure de l'élaboration du questionnaire ainsi que des différents aménagements apportés à la base de donnée (chapitre 4) ; exposé de la méthodologie et du modèle avec d'abord la première méthode utilisée ACM (visant à regrouper les variables pour faire ressortir les différents profils des travailleurs) et ensuite, la régression logistique visant à caractériser les relations entre la variable Retard et les variables indépendantes (chapitre 5) ; modèles économétriques (chapitres 6 et 7) ; étude des différents facteurs expliquant la congestion du trafic routier bruxellois ainsi que des pistes susceptibles d'améliorer la situation (chapitre 8) ; conclusion et identification des limites de l'étude (chapitre 9).

Les données nécessaires à cette étude qualitative ont été collectées via questionnaires. Environ 60% des réponses ont été obtenues en face à face, auprès d'interlocuteurs pris au hasard. Le reste a été récolté via des courriels envoyés au sein de l'entourage. Le but était de sonder un nombre suffisant de personnes utilisant différents moyens de transport, provenant de différentes régions et travaillant dans différents secteurs d'activités. Cette collecte d'informations a permis de constituer une base de données de qualité.

2. La problématique de la mobilité en Belgique

Est-ce que le problème de mobilité mérite vraiment un titre comme celui de cet article "Le chagrin des Belges" ¹ ?

L'expérience personnelle – même limitée à l'écoute des flashs d'info trafic – incite à répondre positivement à cette question et il suffit d'interroger son entourage pour réaliser rapidement qu'un phénomène chronique de retard est bien présent. Qu'ils utilisent la voiture ou les transports en commun, arriver à temps au travail ne semble pas évident pour beaucoup de gens. Et après le travail, ils éprouvent aussi souvent des difficultés pour récupérer à temps leurs enfants à la crèche ou à l'école, leurs rendez-vous médicaux, etc.

Ce sentiment intuitif est confirmé par les chiffres. L'enquête « La frustration des belges » ², commandée par Goodyear, montre qu'une personne devant se déplacer sur quatre éprouve de la frustration. Avec 26% de mécontents, le train subit les conséquences des nombreux retards. Les usagers des bus et des trams se plaignent quant à eux des dessertes trop souvent mal adaptées. Et aussi étonnant que cela puisse paraître, le niveau de frustration des automobilistes est au même niveau que celui des cyclistes (23% d'insatisfaits).

La circulation dans les grandes métropoles belges est un problème majeur. À Bruxelles, la destination principale, les ralentissements débutent traditionnellement dès 6 heures par l'encombrement des autoroutes E19 et E40 venant de Flandre ; suivent la E429 depuis Tournai, la E19 venant de Mons, la E40 venant de Liège et la E411 desservant Namur et le Brabant wallon. Mais il n'y a pas que les autoroutes. Les points névralgiques de notre pays se situent (en longueur de file) à la chaussée de Haacht et à la chaussée de Waterloo (21 km), à Zellik (32 km) et à Grand-Bigard, mais aussi entre Jemeppe-sur-Meuse et Liège (29 km). Les retards peuvent culminer à 91 minutes sur le ring entre Zellik et Woluwe Saint-Étienne (Zaventem). En plus de Bruxelles, trois autres villes belges – Anvers, Charleroi et Gand - figurent au triste palmarès des 25 villes européennes les plus embouteillées. Par comparaison, une seule ville italienne (Milan) y est recensée. Notre pays est donc le plus embouteillé d'Europe, sa capitale trônant en tête du classement pour les cités européennes et même nord-américaines ³.



¹ Dossier mobilité – Le chagrin des Belges - Télépro 4 janvier 2018 par Alice Kriescher

² Source / Infographie : Goodyear 24 mars 2015 : « Le chagrin des Belges ».

³ Rapport annuel Trafic Scorecard de la société américaine Inrix qui répertorie les embouteillages dans le monde entier (<http://inrix.com/>).

Chaque automobiliste belge a perdu 21 heures dans les embouteillages en 2017 ⁴. À Bruxelles on atteint 41,3 heures/an, soit 12% du temps de conduite total ⁵. Le parc automobile belge compte 6 millions de voitures et est en croissance permanente. Des études récentes ⁶ ont montré que d'ici 2030, la croissance en demande de mobilité des personnes sera de 22%, principalement pour les loisirs et celle de la demande de transport de marchandises atteindra 64%.

L'alternative la plus fréquente à la voiture est le train. La SNCB et Infrabel affirment que 90% du trafic voyageurs arrive à l'heure en Belgique ⁷. En apparence, la performance paraît bonne par rapport aux bons élèves européens. Les CFFS (Chemins de fer fédéraux suisses) annoncent aussi que 90% des trains helvétiques ⁸ arrivent à l'heure ⁹. Mais l'association navetteurs.be se montre nettement plus critique : *"Si nous gardons espoir que la régularité s'améliorera, force est de constater que les taux de ponctualité ne rencontrent pas actuellement nos attentes. Quatre convois sur dix sont en retard"* ¹⁰. Les chiffres globaux de ponctualité camouflent probablement des situations bien différentes selon que l'on se trouve aux heures de pointe – avec impact sur les retards - ou aux heures creuses.

Quel est le coût de cette situation ? Sur base de données fournies par des organisations patronales comme Agoria ou Beci, le cabinet du ministre fédéral de l'Économie Kris Peeters (CD&V) a évalué la facture des embouteillages à plus de 100 millions d'euros pour Bruxelles et à plus de 150 millions pour Anvers. En y additionnant les 32 millions d'heures perdues, la pollution de l'air et sonore, les accidents et la perte de productivité, l'usure de l'infrastructure, on arrive à une facture annuelle de 511 millions d'euros pour l'ensemble de l'économie. Selon l'OCDE, la perte économique provoquée par les embouteillages représente entre 1% et 2% du PIB du pays concerné ^{11 12}. Face à ces chiffres, l'on est en droit de s'interroger quant aux mesures prises par les autorités pour diminuer cette lourde perte économique.

Intéressons-nous à présent aux causes du problème. François Janne d'Othée ¹³ et le blogueur Laurent Vermeersch exposent ainsi leur vision « catas-traffic » des raisons de (l'im)mobilité belge.

- La distance importante entre le domicile des travailleurs et le lieu de leur emploi, souvent concentré dans les grandes métropoles.

⁴ <https://www.lecho.be/economie-politique/international/general/bruxelles-une-des-villes-les-plus-embouteillees-d-europe/9864564.html> - Journal L'Echo du 20/2/2017 – source : Rapport annuel Traffic Scorecard de la société américaine Inrix. –page 2 – le 28/4/2018

⁵ <https://www.lecho.be/economie-politique/international/general/bruxelles-une-des-villes-les-plus-embouteillees-d-europe/9864564.html> - Journal L'Echo du 20/2/2017 – source : Rapport annuel Traffic Scorecard de la société américaine Inrix. - page 3 – le 28/4/2018

⁶ Beci : Le livre blanc de la mobilité, Vincent Campeol, chargé de recherches et principal rédacteur du texte.

⁷ Dossier mobilité – Le chagrin des Belges - Télépro 4 janvier 2018 par Alice Kriescher –page102

⁸ Ponctuel comme un train suisse ? | L'interconnexion n'est plus assurée - https://nortonsafe.search.ask.com/web?q=ponctualit%C3%A9%20des%20chemins%20de%20fer%20suisse&o=APN11912&prt=NAV&chn=25700&geo=FR&ver=22.11.2.7&locale=fr_FR&tpr=111&guid=A884B11A-6750-424E-A79E-599B2BCF080C&doi=2016-09-01 – le 5/2/2018

⁹ Pour que la comparaison soit vraiment valable, il faudrait savoir exactement après combien de minutes un retard est comptabilisé dans chaque réseau.

¹⁰ Dossier mobilité – Le chagrin des Belges - Télépro 4 janvier 2018 par Alice Kriescher –page102

¹¹ https://www.rtf.be/info/societe/detail_embouteillages-le-chagrin-des-belges?id=8938753 page 2- le 28/4/2018

¹² <https://www.iweps.be/indicateur-statistique/pib-en-volume/> PIB de la Belgique = 395 833 millions d'euros (2017)

¹³ Bruxelles : Ceci n'est pas une ville, L'âme des peuples François Janne d'Othée.

- Le manque d'infrastructure, de transports en commun efficaces, de parkings de délestage et de pistes cyclables, les trains bondés et leurs retards incessants, les projets de RER sans cesse reportés.
- Les avantages fiscaux en matière de voitures de société, lesquelles représentent 40% du parc automobile belge et près de 60% des files du matin.
- Le flot de camions transitant par la Belgique.
- Un réseau autoroutier trop centré sur Bruxelles, menant inévitablement à un engorgement du Ring 0.

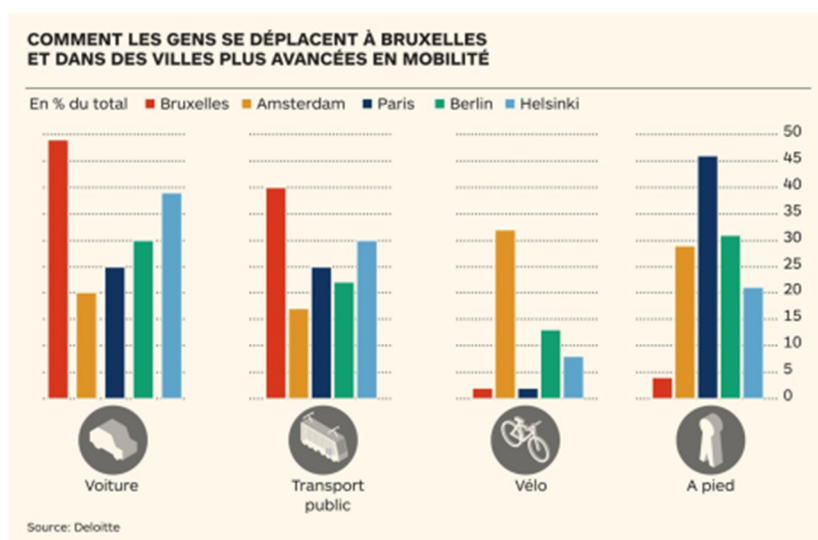
Pour aller plus loin, un entretien a été mené avec Eric Van Rompuy, élu CD&V, président de la commission des finances et du budget au Parlement flamand. Fort de son expérience d'échevin chargé des finances, de l'aménagement du territoire et de la planification urbaine de Zaventem, cet homme politique est bien au fait des problèmes de mobilité à Bruxelles. Il en attribue les causes :

- Au fait que plusieurs grandes entreprises, institutions publiques et instances internationales (UE, OTAN) ont établi leur siège à Bruxelles ;
- à la proximité de l'aéroport de Zaventem qui voit défiler pas moins de 21 millions de passagers dont 80% utilisent la voiture pour seulement 11,8% les métro-trams-bus et 7,4% les chemins de fer ;
- à la structure du réseau routier belge qui oblige les transitaires, d'où qu'ils viennent et où qu'ils aillent, à passer par le ring de Bruxelles ;
- Aux nombreux et interminables chantiers ;
- à l'augmentation incessante du nombre des voitures de société. En 10 ans, leur nombre a augmenté de 55%, entraînant un manque à gagner fiscal de près de 2 milliards pour l'État.

Eric Van Rompuy affirme que dans un pays à forte densité de population, les solutions sont délicates à mettre en place. Depuis 2004, de nombreuses idées de désengorgement de Bruxelles ont été exprimées : un projet pour optimiser le ring de Bruxelles en l'élargissant de 4 à 6 bandes, voire 8 à certains endroits ; un troisième proposant une déviation pour les navetteurs venant de Louvain vers Bruxelles, un autre encore suggérant d'augmenter les lignes de métro et notamment une liaison qui connecterait directement l'aéroport de Zaventem au centre économique de Bruxelles. Malheureusement, aucun de ces projets n'a abouti jusqu'à présent. La complexité du paysage institutionnel belge et notamment, l'éclatement des compétences en matière de mobilité, ne facilite pas le dégagement de compromis. Monsieur Van Rompuy attribue également la persistance des problèmes au fait que le politique a souvent tendance à chercher des solutions provisoires, sans développer une vision sur le long terme. Il n'est pas vraiment optimiste quant à la résolution du problème de la mobilité.

Pour éviter de tomber dans ce fatalisme, il y a tout de même des pistes porteuses d'espoir.

Comme on le voit dans le graphique ci-après, le Belge adore particulièrement utiliser sa voiture, tant pour ses loisirs que pour ses déplacements professionnels.



Il s'agit là d'habitudes de comportement. Il devrait être possible de les changer via des campagnes de « rééducation » pour inciter les Belges à éviter de prendre systématiquement sa voiture pour des petits trajets ou recourir davantage au covoiturage.

Des solutions simples et aisées à mettre en place permettraient également de stimuler l'attrait des transports alternatifs. Prenons l'exemple de la création d'un titre de transport unique, similaire à la "Oyster Card" londonienne, qui donnerait accès à tous les moyens de transports disponibles : trains, trams, métros, bus, vélos et voiture partagée pour les zones non desservies par les transports publics.

De plus en plus, le recours aux solutions offertes par l'économie de partage devrait aussi permettre d'améliorer la mobilité tout en réduisant la facture économique-sociale. Une étude de Deloitte ¹⁴ évalue les économies de coût possibles selon les modes de déplacement. Ainsi, le covoiturage permettrait 35% d'économie, la voiture autonome jusqu'à 50% et le covoiturage en voiture autonome réduirait la facture à un tiers du montant actuel.

L'économie se développe de plus en plus dans le secteur des services, c'est-à-dire là où les possibilités de télétravail sont les plus étendues. Les moyens technologiques actuellement à la disposition des entreprises autoriseront probablement de plus en plus de gens de travailler chez eux.

Diverses études prévoient qu'à l'aube de 2030 et à l'échelle mondiale, 50% des véhicules seront électriques, 35% des déplacements se feront à bord de voitures partagées dont 10% des km parcourus à bord de voitures autonomes. Un tiers des déplacements sera effectué via des moyens de transport qui n'existent pas encore aujourd'hui, comme les bus sans chauffeur ou les drones.

Terminons en citant San Sluismans, expert en mobilité chez Deloitte : "Il ne faut pas attendre passivement que les choses changent. Nous sommes arrivés au point d'inflexion. C'est le moment de montrer l'exemple."

¹⁴ Deloitte 2018 étude "Future of Mobility".

3. Revue de la littérature

Les différents ouvrages et articles consultés traitent de la mobilité dans son ensemble : les différents moyens de déplacement, les motifs de déplacement, les méthodologies pour collecter les données et éventuellement les conséquences des moyens de transports.

Hubert et Toint (1999) ont étudié ¹⁵ l'organisation des activités quotidiennes des individus ainsi que les détails de leurs déplacements vers ces activités. Pour leur analyse, ils procèdent par une enquête en plusieurs étapes : cibler la population, cadrer son échantillon et passer à la sélection. Le choix de la formalisation de cette enquête est minutieusement réfléchi. Ils ont fait appel à une société de sondage qui a envoyé un questionnaire-papier aux ménages en utilisant le Registre national en tant que source. Les avantages de la motorisation, son importance et son utilité ressortent de l'étude mais également le fait qu'elle est la cause de la dégradation des conditions de circulation. L'augmentation de la motorisation évolue de manière plus rapide que la capacité des infrastructures pouvant l'accueillir. Il en résulte une asphyxie. Le choix de l'usage d'un véhicule privé ou des transports en commun a de multiples raisons. Qu'elles soient personnelles, économiques et/ ou logistique n'aura pas d'incidence directe sur la situation. Seuls les changements de comportement et d'habitudes pourrions l'améliorer.

Pour Lebrun et Al (2012), les résultats ¹⁶ sont similaires à ceux de Hubert et Toint (1999). Les données ont été collectées auprès d'un échantillon aléatoire de ménages résidant en Belgique par le biais d'un questionnaire concernant le ménage dans son ensemble et d'un autre à remplir par chaque membre de la famille de plus de six ans. Il ressort de cette étude, qui est une comparaison entre l'étude MOBEL ¹⁷ (1999) et l'étude BELDAM ¹⁸ (2010), une similitude des résultats obtenus. Toutefois, on remarque une augmentation de l'éloignement domicile-travail, de la densité de circulation et du temps consacré au déplacement. Il y a statu quo de certains paramètres tels que les motifs du déplacement, l'augmentation de la densité de circulation aux heures de pointe et les obligations administratives.

En ce qui concerne l'enquête de Bouéroux et al (2010), l'objectif de la recherche ¹⁹ est la détermination des liens entre l'éloignement géographique domicile-travail et les retards. Ils expliquent qu'il faut également tenir compte d'un ensemble de risques et de perturbations, tant au niveau professionnel qu'au niveau familial, ainsi que des réponses apportées par les entreprises. Les auteurs ont collecté leurs informations via des entretiens semi-directifs avec un échantillon de travailleurs ainsi qu'une grille d'entretien codée, réalisée auprès des représentants du personnel et de la direction des ressources humaines.

¹⁵ <http://pun.be/fr/livre/?GCOI=99993102906770> – La mobilité quotidienne des Belges –le 5-5-2018

¹⁶ Cahiers de l'Observatoire de la mobilité de la Région de Bruxelles-Capitale https://mobilite-mobiliteit.brussels/sites/default/files/les_pratiques_de_deplacement_a_bruxelles_analyses_appfondies.pdf - le 5-5-2018

¹⁷ <https://mobilit.belgium.be/sites/default/files/downloads/mobel1999.pdf> - Mobilité Quotidienne en Belgique –le 5-5-2018

¹⁸ https://mobilit.belgium.be/sites/default/files/downloads/enquete_resultats.pdf - Beldam -le 5-5-2018

¹⁹ Conditions de transport & santé des salariés et des entreprises – Technologia – janvier 2010 <http://www.technologia.fr/blog/wp-content/uploads/2010/03/Enquête-Transports-Technologia.pdf> le 5-5-2018

L'enquête montre que la multiplication des moyens de transport et/ou les mauvaises conditions de ceux-ci augmentent le risque de retard, lequel peut à son tour influencer négativement le moral et le rendement du salarié. Si l'entreprise se situe dans une zone péri-urbaine, les salariés se rendront plus volontiers en voiture qu'avec les transports en commun. Ceux-ci sont en effet peu disponibles et leurs horaires pas toujours adaptés. Par ailleurs, le choix géographique du domicile dépend fortement du prix des loyers et de l'immobilier et du revenu des salariés. Toutefois, l'utilisateur d'un véhicule automobile individuel ou du covoiturage sera confronté à divers risques comme un accident ou des embouteillages, ce qui peut rallonger le temps de trajet.

L'étude fait apparaître une forte corrélation entre le temps du trajet et la distance parcourue. Une première analyse montre que les conditions dans lesquelles se déroule le trajet domicile-travail a un impact important sur la santé des salariés comme sur l'organisation de l'entreprise. L'horaire variable permet à beaucoup gens de trouver un meilleur équilibre professionnel et familial susceptible d'avoir un impact positif sur leur rendement. Plus une journée de travail est longue (prestation et trajet), plus l'équilibre vie professionnelle-vie familiale est perturbé. Les pistes permettant un meilleur équilibre sont le télétravail et l'horaire.

La météo peut aussi être une cause de retard comme l'illustre l'expérience de Simon Demeulemeester²⁰ (2017), ancien rédacteur en chef de Knack.be/LeVif.be.

En résumé, les deux premiers travaux traitent essentiellement de la mobilité en Belgique et à Bruxelles, de l'influence sur la distance et le temps de trajet. Le troisième article traite du choix forcé des modes de déplacement par manque d'alternatives et de l'impact de ce choix sur le travailleur, sur son rendement et les conséquences néfastes pour sa santé physique et psychologique. Ceci s'est révélé utile pour l'élaboration du questionnaire. Le dernier article met en évidence l'influence de la météo sur un aspect de la mobilité. Il est également à l'origine d'un certain nombre de questions relatives à la météo dans le questionnaire.

Aucun de ces auteurs n'a fait le lien entre problèmes de mobilité et retard. Leurs travaux ne permettent donc pas d'affirmer ou d'infirmer que les retards sont une conséquence d'une mobilité désastreuse. La présente étude apporte donc un complément intéressant aux travaux existants.

²⁰ <http://www.levif.be/actualite/belgique/en-train-ou-en-voiture-etre-a-l-heure-au-travail-est-devenu-mission-impossible/article-opinion-600415.html> -le 5-5-2018

4. Constitution de la base de données

4. 1 Élaboration d'un questionnaire

Pour la constitution de la base de données, un questionnaire spécifique comportant une quinzaine de questions a été conçu. Pour une mesure plus fine des variables principales, elles ont été subdivisées en sous-questions. Certaines questions secondaires de contrôle ont été bornées, c'est-à-dire que la personne interrogée avait le choix entre plusieurs plages de valeurs numériques. Ceci a permis de dissiper les flous et de réduire la subjectivité.

4. 2 Source des données

Les données ont été collectées à Bruxelles, Namur, Liège et Arlon. Sur les 339 personnes contactées, 210 questionnaires se sont révélés utilisables (voir annexe 1) ²¹. Les individus interrogés travaillent dans les 3 régions du pays mais aussi au Luxembourg. Ces données, en coupe transversale, permettent d'observer l'ensemble des comportements de chaque travailleur ²² par rapport à son mode de déplacement. Il existe deux types de variables : quantitatives et qualitatives. Les variables binaires sont considérées comme qualitatives. La variable expliquée est "Retard". Toutes les autres variables sont explicatives (voir annexe 3) ²³.

4. 3 Aménagement de la base de données

Quelques adaptations simplificatrices ont été effectuées. Pour "Moyen de transport", les codes de départ 1; 0,5; 0,25, inutilisables en tant que tels dans le traitement des informations, ont été remplacés uniquement par le moyen prioritaire utilisé en le cotant "1". Mais pour ne pas perdre de l'information, une nouvelle variable "Utilisation d'au moins deux moyens de transport" a été créée. "Auto" et "Train", variables très importantes, ont été gardées seules. "Bus + Tram + Métro" ont été fusionnées en une seule variable "Transports en commun". "A pied + Vélo" ont été incorporées dans une seule variable : "Mobilité douce".

Les données qualitatives ordonnées "Raison du mode de transport" et "Type de problème rencontré" ont été condensées. Du classement initialement prévu en 6 classes (0, 1, 2, 3, 4, 5), nous n'avons retenu que 0, 1 et 5. La cotation 1 a absorbé la cotation 2 et la cotation 5 a absorbé les cotations 3 et 4.

Pour pouvoir traiter le problème en Analyse des Correspondances Multiples (ACM), (voir paragraphe 5), les variables ont été rendues homogènes en transformant les variables initialement numériques en variables qualitatives ²⁴. La variable expliquée "Retard" a été remplacée par une variable binaire qui ne précise plus une valeur mais indique simplement s'il y a un retard ou pas (0 ou 1), ce qui répond mieux à la problématique. Les variables quantitatives "Distance parcourue" et "Marge de Sécurité" ont été découpées en classes. Chaque classe a alors pu être remplacée par une modalité. Le nombre suffisant de modalités par classe est de quatre ²⁵. Exemple : Distance 0 à 5 km = "1". La variable "Population du lieu de travail" n'a pas été prise en considération.

²¹ Un tableau détaillé des sources de données est présenté à l'annexe 1.

²² Voir Annexe 4 – Remarques sur les Horaires mobiles, la Direction et les Cadres.

²³ Chaque variable et leurs instances correspondantes sont présentées et commentées à l'annexe 2.

²⁴ Escofier, Brigitte et Pagès, Jérôme. 2016 - 5ème édition. Analyses factorielles simples et multiples. Paris : Dunod, 2016 - 5ème édition. -p. 105

²⁵ Escofier, Brigitte et Pagès, Jérôme. 2016 - 5ème édition. Analyses factorielles simples et multiples. Paris : Dunod, 2016 - 5ème édition. -p. 107 - "Dans la pratique, les variables qualitatives étudiées en ACM résultent souvent d'une transformation de variables numériques... On peut vouloir rendre homogènes des données qui se composent initialement de variables numériques et de variables qualitatives. Le fait de transformer les variables numériques en variables qualitatives permet de traiter l'ensemble de ces variables par l'ACM".

5. Présentation de la méthodologie et du modèle

Pour traiter notre base de données, il a fallu identifier les méthodes adéquates. Deux méthodes ont été jugées nécessaires.

5. 1 Étape 1–Analyse des Correspondances Multiples (ACM)

Pour traiter ce type de bases de données, la méthode de l'ACM ²⁶ a été retenue. Elle permet d'étudier une population de I individus décrits par J variables qualitatives. "En effet l'application la plus courante pour l'ACM est le traitement de l'ensemble des réponses à une enquête" ²⁷, ce qui est le cas ici. Chaque question correspond à une variable dont les modalités sont les réponses proposées (parmi lesquelles chaque répondant doit faire un choix unique).

Mais deux difficultés se sont posées. Si la base de données était essentiellement composée de variables qualitatives, quelques variables quantitatives étaient néanmoins présentes. Pour pouvoir appliquer la méthode ACM, les données ont donc été rendues homogènes via le codage des variables numériques par classe (4 classes sont suffisantes) ²⁸, en découpant des intervalles de variation. En second lieu, si cette méthode permet d'identifier les variables qui ont un lien direct avec la mobilité, elle ne permet pas une interprétation simple et claire des résultats. D'où la décision d'utiliser cette première méthode simplement comme filtre et de ne retenir qu'un sous-ensemble de données qui ont été traitées ensuite par une deuxième méthode.

5. 2 Étape 2 – La Régression Logistique ²⁹

La pratique de la régression logistique est très proche de celle de la régression linéaire. La régression linéaire permet de caractériser les liens entre une variable à expliquer (Y) quantitative et des variables explicatives ($X_1, X_2, X_3, \dots, X_n$) au moyen du modèle présenté par la formule suivante :

$$Y = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \beta_3 * X_3 + \dots + \beta_n * X_n + \varepsilon$$

Lorsque la variable dépendante Y est binaire, le modèle de régression linéaire n'est pas des plus adaptés ³⁰. Il est donc nécessaire d'utiliser un modèle plus ajusté permettant de relier les variables explicatives à la variable qualitative (Y) à prédire. L'astuce de la régression logistique consiste non pas à modéliser la variable qualitative Y mais la probabilité que la variable dépendante prenne les valeurs 1 ou 0.

²⁶ Les principes de cette méthode peuvent remonter à Guttman (1941) et à Burt (1950) - Le nom d'Analyse des Correspondances Multiples figure pour la première fois dans Lebart (1975)

²⁷ Escofier, Brigitte et Pagès, Jérôme. 2016 - 5ème édition. Analyses factorielles simples et multiples. Paris : Dunod, 2016 - 5ème édition. -p. 89

²⁸ Escofier, Brigitte et Pagès, Jérôme. 2016 - 5ème édition. Analyses factorielles simples et multiples. Paris : Dunod, 2016 - 5ème édition. -p. 107

²⁹ <http://www.em-consulte.com/en/article/842576> - le 12-5-2108

³⁰ Remarque : dans un modèle de régression linéaire, il peut y avoir des variables explicatives binaires.

Le modèle logistique permet une expression non linéaire, variant entre 0 et 1. La régression logistique repose sur l'hypothèse fondamentale suivante :

$$\text{Evidence} = \text{Ev}(p) = \ln p / (1 - p)^{31} = \text{LOGIT} = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \beta_3 * X_3 + \dots \beta_n * X_n + \varepsilon$$

où p est la distribution conditionnelle des X sachant la valeur prise par Y. Par transformation logarithmique on obtient :

$$P(Y) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \dots \beta_n * X_n)}}$$

P(y) est la probabilité que Y arrive. Lorsque la valeur prédite est supérieure à 0,5, l'événement est susceptible de se produire, alors que lorsque cette valeur est inférieure à 0,5, il ne l'est pas ³².

³¹ Mesure nommée "évidence" popularisée par L. J. Good -https://fr.wikipedia.org/wiki/R%C3%A9gression_logistique - page 2/8

³² <http://spss.espaceweb.usherbrooke.ca/pages/stat-inferentielles/regression-logistique.php> - le 11/6/2018

6. Application de la méthode ACM

6. 1 Les hypothèses

H1 : les moyens de transport et les types de problèmes rencontrés sont une cause du retard.

H2 : la marge de sécurité, la région d'habitation et le lieu de travail augmentent la probabilité d'arriver en retard.

H3 : le retard dépend des caractéristiques des travailleurs.

6. 2 Choix des variables

Après le nettoyage des données ³³, toutes les variables restantes ont été considérées comme des variables actives afin de pouvoir répondre aux trois hypothèses de départ.

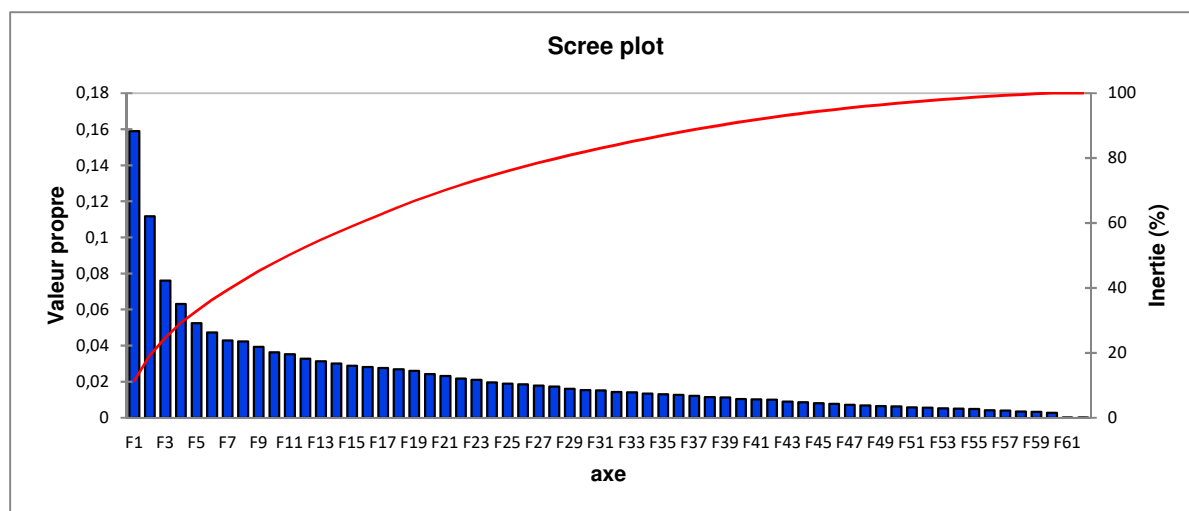
6. 3 Choix du plan factoriel

La projection de toutes les variables permet de constituer l'histogramme (voir graphique 1) en choisissant les axes factoriels F1, F2 et F3 qui représentent 24.70% de l'inertie totale du nuage. Cela signifie que les variables représentent 24.70% de la réalité sur le plan factoriel. Ce résultat nous permet de définir des différents profils qui expliqueront la variable retard. Cette représentation moyenne (24.70%) ne pose pas de réel problème pour l'analyse en ACM.

Tableau n° 1 Les valeurs propres et le pourcentage d'inertie

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
Valeur propre	0,159	0,112	0,076	0,063	0,052	0,047	0,043	0,042	0,039	0,036
Variabilité (%)	11,323	7,958	5,415	4,494	3,735	3,372	3,048	3,018	2,801	2,583
% cumulé	11,323	19,281	24,696	29,190	32,925	36,297	39,345	42,363	45,164	47,747

Graphique n°1 Histogramme des valeurs propres en fonction des axes



³³ Voir annexe 5

Cette méthode a l'avantage de regrouper les variables (donc d'en diminuer le nombre) et de faire ressortir les différents profils d'individus. Le graphique qui est la projection dans le plan F1/F3 (voir annexe 3) devrait permettre de repérer les variables les plus pertinentes pour continuer l'étude. Cependant, on voit qu'en vertu de son principe même, cette méthode n'est pas très pratique. En effet, comment lire facilement un graphique surchargé de tant d'informations ? Seuls les quadrants F1+/F3+ et F1-/F3- livrent des résultats exploitables. Plus les points sont éloignés du centre, plus on augmente les chances de retard. Le quadrant -/- apporte une contribution négative. D'où la proposition du tableau 2, plus compréhensible via le catalogage des caractéristiques des axes vectoriels.

6. 4 Caractéristiques des axes factoriels

Le tableau 2 est très important car il pallie l'imprécision du graphique de l'annexe 3 et nous aidera à identifier les profils intéressants pour cette étude. Seules les variables avec une contribution supérieure à l'inertie totale de 1,404 seront retenues.

Tableau n °2– Caractéristiques des axes vectoriels

AXE F1					
Positif			Négatif		
Variable	Inertie	Cos ²	Variable	Inertie	Cos ²
pas le choix 0	0,017	0,192	pas le choix 1	0,028	0,278
par habitude 0	0,043	0,497	par habitude 5	0,020	0,362
économique 0	0,048	0,606	sécurisant 5	0,029	0,419
fiable ponctuel 0	0,037	0,410	aucun problème 1	0,026	0,296
rapide 0	0,046	0,459	stationnement 1	0,031	0,372
sécurisant 0	0,046	0,568	fiabilité matériel 1	0,024	0,281
aucun problème 0	0,021	0,296			
stationnement 0	0,028	0,401			
incident 0	0,037	0,384			
fiabilité matériel 0	0,036	0,500			
fiabilité horaires 1	0,018	0,176			
fiabilité horaires 0	0,036	0,476			
AXE F2					
Variable	Inertie	Cos ²	Variable	Inertie	Cos ²
écologique 1	0,034	0,281			
écologique 5	0,041	0,299	être autonome 0	0,028	0,195
économique 1	0,021	0,144	être autonome 1	0,034	0,174
fiable ponctuel 1	0,031	0,190	être autonome 5	0,031	0,017
rapide 5	0,014	0,200	bouchon 0	0,022	0,152
rapide 1	0,031	0,182	bouchon 5	0,020	0,283
au moins deux moyens de transport	0,019	0,251	fiabilité matériel 5	0,017	0,109
			fiabilité horaires 5	0,061	0,421
			auto 0	0,095	0,731
			auto 1	0,045	0,731
			train 0	0,014	0,481
			train 1	0,077	0,481
			transport en commun 1	0,016	0,094
			au moins deux moyens de transport	0,029	0,251
AXE F3					
Variable	Inertie	Cos ²	Variable	Inertie	Cos ²
fiable ponctuel 5	0,020	0,148	horaires mobile 0	0,015	0,080
sécurisant 1	0,027	0,105	retard 1	0,018	0,151
aucun problème 5	0,051	0,227	Bruxelles travail 1	0,041	0,193
bouchon 1	0,019	0,080	Wallonie travail 0	0,063	0,396
incident 1	0,020	0,090	distance 4	0,018	0,074
manuel 1	0,021	0,104	Flandre habitation 1	0,037	0,013
retard 0	0,025	0,151			
mobilité douce 1	0,050	0,151			
Wallonie travail 0	0,063	0,396			
Wallonie travail 1	0,048	0,396			
mobilité douce 1	0,050	0,195			
distance parcourue 1	0,035	0,147			
Wallonie habitation 1	0,018	0,195			

6. 5 Interprétation

Remarquons que l'axe 2 ne s'affiche pas sur le graphique et que ses données sont incluses dans les axes 1 et 3 d'où les profils suivants sont ressortis. Seuls les quadrants F1+/F3+ et F1-/F3- seront pris en considération.

Profil 1 (négatif)

Dans le quadrant -/-, on discerne des salariés susceptibles d'arriver en retard et dont le profil se décline de la façon suivante : ils habitent loin de leur lieu de travail, à environ 50 km de distance. Généralement, leur occupation est soit à Bruxelles, soit en Flandre. Ils utilisent principalement le train qu'ils trouvent économique. Éventuellement, ils ont recours à un deuxième moyen de déplacement qui peut être l'auto, le bus ou la marche ; lorsqu'ils utilisent la voiture, ils sont souvent victimes des bouchons. Leur souci majeur est l'incertitude des horaires de train.

Profil 2 (positif)

Dans le quadrant +/+, on distingue des salariés susceptibles de ne pas arriver en retard et dont le profil se définit comme suit : ils travaillent en Wallonie et parcourent moins de 5 km pour rejoindre leur lieu de travail. Ils utilisent les transports en commun ou la mobilité douce – marche ou vélo - car il l'a juge fiable et ponctuelle. Ils ne rencontrent généralement aucun problème au cours de leur trajet. Ils ne craignent ni les bouchons ni les incidents durant leur parcours.

6. 6 Résultats

On peut voir sur le graphique de l'ACM qu'il existe un lien significatif entre la variable dépendante Y et les moyens de transport. Ce qui entraîne que l'hypothèse 1 soit vérifiée.

La seconde hypothèse est également validée dans la mesure où le tableau des contributions montre la relation entre la variable Y, les régions et les distances parcourues par leurs inerties élevées : Wallonie T (4,8%), distance parcourue (3,5%), Bruxelles T (4,1%), Flandre H (3,7%), Wallonie H (1,8%), etc.

Les caractéristiques des travailleurs n'ont pratiquement pas d'influence sur les retards. Le fait d'être un homme ou une femme, d'avoir un CDD ou un CDI, ou encore d'être cadre ou ouvrier n'a aucun effet sur les retards éventuels. Ainsi, la troisième hypothèse n'est pas vérifiée.

7. Application de la Régression logistique

7.1 Les hypothèses

$H_0 : \beta_i = 0$ tous les coefficients sont nuls.

$H_1 : \beta_i \neq 0$ il existe au moins un coefficient non nul.

7.2 Choix des facteurs à analyser dans la régression logistique

Ne pouvant pas entrer toutes les variables utilisées dans l'ACM, le choix s'est fixé sur les variables directes. Le logiciel XLSTAT a été utilisé pour cette étude.

Tableau 3 -Les variables retenues pour la régression logistique après simulation ACM

Source	Valeur	Erreur standard	Khi ² de Wald	Pr > Khi ²	Wald Borne inf. (95%)	Wald Borne sup. (95%)	Odds ratio	Odds ratio Borne inf. (95%)	Odds ratio Borne sup. (95%)
Constante	-1,160	0,699	2,754	0,097	-2,530	0,210			
Auto-1	1,448	0,696	4,326	0,038	0,084	2,813	4,256	1,087	16,659
Train-1	2,206	0,904	5,963	0,015	0,436	3,977	9,084	1,546	53,379
Transport en commun-1	1,134	0,854	1,763	0,184	-0,540	2,808	3,108	0,583	16,574
Au moins 2 moyens de transport-1	0,303	0,404	0,562	0,453	-0,489	1,095	1,354	0,613	2,990
Marge de sécurité / min-2	-0,654	0,349	3,521	0,061	-1,338	0,029	0,520	0,262	1,030
Marge de sécurité / min-3	-1,114	0,549	4,117	0,042	-2,191	-0,038	0,328	0,112	0,963
Marge de sécurité / min-4	-2,535	2,383	1,131	0,288	-7,206	2,137	0,079	0,001	8,470
aucun problème-1	0,065	0,434	0,022	0,882	-0,786	0,915	1,067	0,456	2,497
aucun problème-5	-0,421	0,487	0,748	0,387	-1,374	0,533	0,657	0,253	1,704
bouchon-1	0,847	0,630	1,805	0,179	-0,389	2,082	2,332	0,678	8,024
bouchon-5	0,838	0,482	3,024	0,082	-0,107	1,783	2,312	0,899	5,949
stationnement-1	0,068	0,457	0,022	0,882	-0,827	0,962	1,070	0,437	2,618
stationnement-5	0,158	0,497	0,102	0,750	-0,816	1,133	1,172	0,442	3,104
incidents-1	-1,101	0,522	4,446	0,035	-2,124	-0,078	0,333	0,120	0,925
incidents-5	-0,366	0,434	0,712	0,399	-1,216	0,484	0,694	0,296	1,623
fiabilité matériel-1	0,282	0,494	0,326	0,568	-0,686	1,250	1,326	0,504	3,490
fiabilité matériel-5	0,008	0,500	0,000	0,988	-0,973	0,988	1,008	0,378	2,687
fiabilité horaires-1	-0,312	0,501	0,387	0,534	-1,294	0,670	0,732	0,274	1,955
fiabilité horaires-5	0,302	0,625	0,233	0,629	-0,923	1,526	1,352	0,397	4,600

À la lecture des résultats du tableau 3, on peut constater que les variables Transport en commun-1, Au moins 2 moyens de transport-1, Aucun problème-1, Stationnement-1, Fiabilité matériel-5, Fiabilité horaires-5 ont une probabilité critique respectivement de 0,184, 0,453, 0,882, 0,882, 0,988 et 0,629, toutes supérieures à 0,05. Elles ne sont donc pas significatives.

On procède alors à une seconde estimation, en ne retenant cette fois que les facteurs dont l'analyse a montré une relation significative avec $p < 0,2$. Bien que non significatives, on les prendra quand même pour la seconde estimation.

Tableau 4 -Les facteurs retenus pour la régression logistique finale

Source	Valeur	Erreur standard	Khi ² de Wald	Pr > Khi ²	Wald Borne inf. (95%)	Wald Borne sup. (95%)	Odds ratio	Odds ratio Borne inf. (95%)	Odds ratio Borne sup. (95%)
Constante	-1,366	0,686	3,963	0,047	-2,710	-0,021			
Auto-1	1,446	0,684	4,466	0,035	0,105	2,787	4,246	1,111	16,232
Train-1	2,856	0,788	13,139	0,000	1,312	4,401	17,395	3,713	81,496
Transport en commun-1	1,510	0,809	3,487	0,062	-0,075	3,095	4,526	0,928	22,079
Marge de sécurité / min-2	-0,625	0,336	3,462	0,063	-1,284	0,033	0,535	0,277	1,034
Marge de sécurité / min-3	-0,996	0,528	3,553	0,059	-2,031	0,040	0,369	0,131	1,040
Marge de sécurité / min-4	-2,308	2,347	0,967	0,325	-6,908	2,292	0,099	0,001	9,896
bouchon-1	0,886	0,563	2,480	0,115	-0,217	1,989	2,426	0,805	7,311
bouchon-5	1,123	0,428	6,900	0,009	0,285	1,961	3,074	1,330	7,106
incidents-1	-1,067	0,464	5,279	0,022	-1,977	-0,157	0,344	0,138	0,855
incidents-5	-0,230	0,382	0,360	0,548	-0,979	0,520	0,795	0,376	1,682

7.3 Interprétation statistique

Statistiques descriptives

Tableau 5 – Ratio retard/non retard

Variable binaire	Modalités	Effectif	Pourcentage
Pas de retard	0	89	42,4 %
Avec retard	1	121	57,6 %
		210	

On constate que 42,4% des navetteurs arrivent à l'heure et 57,6% en retard. (voir tableau 5).

Coefficients d'ajustement

Tableau 6 – Les coefficients d'ajustement

Statistique	Indépendant	Complet
Observations	210	210
Somme des poids	210,000	210,000
DDL	209	199
-2 Log(Vraisemblance)	286,227	211,736
R ² (McFadden)	0,000	0,260
R ² (Cox and Snell)	0,000	0,299
R ² (Nagelkerke)	0,000	0,401

La question est de savoir si notre modèle est proche de la réalité et permet de prédire le comportement de la variable indépendante Y. Le pseudo R² se calcule via diverses méthodes. Parmi elles, nous retrouvons la méthode de Cox Snell (R² = 0,299) ou encore de Nagelkerke (R² = 0,401) pour les plus connues. Toutefois il est reconnu que le modèle de McFadden (R² = 0,260) est le plus performant et le plus adapté à la régression logistique³⁴. Comme la valeur de ce coefficient peut varier entre 1 = très bon modèle et 0 = modèle trivial, la qualité de notre modèle peut être qualifiée assez bonne.

Test de l'hypothèse nulle H₀

Afin de tester l'hypothèse que le modèle améliore statistiquement la compréhension du phénomène observé, on va comparer un modèle naïf, c'est-à-dire sans variables explicatives, au

³⁴ <https://wikilean.com/articles-analyse-regression-regression-logistique/> pages 8 et 10 –le 11/6/2018

modèle proposé. Si celui-ci améliore statistiquement la compréhension du phénomène étudié, alors l'hypothèse nulle d'égalité statistique des deux modèles sera rejetée.

$Y = 0,576$ (variable retard)

Tableau 7 - Test de l'hypothèse nulle

Statistique	DDL	Khi ²	Pr > Khi ²
-2 Log(Vraisemblance)	10	74,491	< 0,0001

La statistique de la Log vraisemblance est égale à 211,736 que l'on compare à un χ^2 lu dans une table avec un seuil de 95% et 10 degrés de liberté. Le χ^2 (10) vaut 74,491 < 211,736 ce qui conduit au rejet de l'hypothèse H_0 .

Test de type II

Tableau 8- Test de type II

Source	DDL	Khi ² (Wald)	Pr > Wald	Khi ² (LR)	Pr > LR
Auto	1	4,466	0,035	44,583	< 0,0001
Train	1	13,139	0,000	56,948	< 0,0001
Transport en commun	1	3,487	0,062	43,121	< 0,0001
Marge de sécurité / min	3	5,819	0,121	46,012	< 0,0001
bouchon	2	7,017	0,030	46,517	< 0,0001
incidents	2	6,112	0,047	45,409	< 0,0001

Les tests de type II correspondent à l'association globale entre variables explicatives et la variable expliquée. Il rassemble toutes les variables significatives qui expliquent la variable dépendante.

Test de Hosmer-Lemeshow

Tableau 9 - Test de Hosmer-Lemeshow

Statistique	Khi ²	DDL	Pr > Khi ²
Statistique de Hosmer-Lemeshow	9,708	7	0,206

Il s'agit d'un test d'ajustement de la régression logistique qui indique dans quelle mesure nos données correspondent à celles du modèle. Plus précisément, si la valeur p est < 5%, cela signifie que le modèle ne convient pas. Or ici $p = 0,206 > 0,05$, ce qui permet de conclure que le modèle est bien ajusté.

Interprétation des paramètres du modèle

Pour interpréter ces résultats (voir tableau 3), on observera le signe des coefficients β qui indique si la variable a une influence positive ou négative sur la probabilité d'arrivée en retard p. On appréciera ensuite la significativité des coefficients à l'aide des odds ratios^{35 36}.

³⁵ L'odds ratio ou "rapport de cote" correspond au nombre de fois d'appartenance à un groupe lorsque la valeur du prédicteur augmente de 1. Il associe la variable explicative à la variable expliquée. Odd ratio est l'exponentiel du coefficient β .

³⁶ <https://help.xlstat.com/customer/fr/portal/questions/16253648-odds-ratio> page 1 – 18/6/2018 – Pour obtenir les odds-ratio, il suffit de prendre l'exponentielle des coefficients. En effet, les coefficients sont exprimés en log (odds-ratio). Exemple : $\exp 1,448 = 4,256$. L'Intervalle de Confiance IC à 95 % correspond aux bornes supérieure et inférieure de Wald.

Auto : lien significatif ; $p(0,035)$ est $< 0,05$; odds ratio est de 4,246; l'intervalle de confiance exclut la valeur 1 ; le signe du coefficient est positif.

Conclusion : il y a un risque multiplié par 4,246 d'arriver en retard en utilisant l'auto comme moyen de transport pour se rendre à son lieu de travail.

Train : lien significatif ; $p(0,000)$ est $< 0,001$; odds ratio est de 17,395; l'intervalle de confiance exclut la valeur 1 ; le signe du coefficient est positif.

Conclusion : l'utilisation du train comme moyen de transport multiplie le risque de retard par 17,395.

Transports en commun : lien significatif ; $p(0,062)$ est $< 0,1$; odds ratio est de 4,526; l'intervalle de confiance exclut la valeur 1 ; le signe du coefficient est positif.

Conclusion : l'utilisation des transports en commun comme moyen de transport multiplie le risque de retard par 4,526.

Marge de sécurité : lien faiblement significatif ; $p(0,063 \text{ et } 0,059)$ est $< 0,1$; odds ratio est de 0,059 ; l'intervalle de confiance exclut la valeur 1 ; le signe du coefficient est négatif.

Conclusion : prendre une marge de sécurité au départ diminue le risque de retard respectivement par 0,63 et 0,59.

Bouchon : lien significatif ; $p(0,09)$ est $< 0,05$; odds ratio est de 3,074; l'intervalle de confiance exclut la valeur 1 ; le signe du coefficient est positif.

Conclusion : un bouchon multiplie le risque de retard par 3,074.

Incident 1 : lien significatif ; p est $< 0,05$; odds ratio est de 0,344; l'intervalle de confiance exclut la valeur 1 ; le signe du coefficient négatif.

Conclusion : les travailleurs qui estiment ne pas rencontrer d'incident au cours de leur trajet voient diminuer leur risque de retard par 0,344.

Tableau de classification pour l'échantillonnage d'apprentissage

Tableau 10 - Tableau de classification

de \ Vers	0	1	Total	% correct
0	46	43	89	51,69%
1	19	102	121	84,30%
Total	65	145	210	70,48%

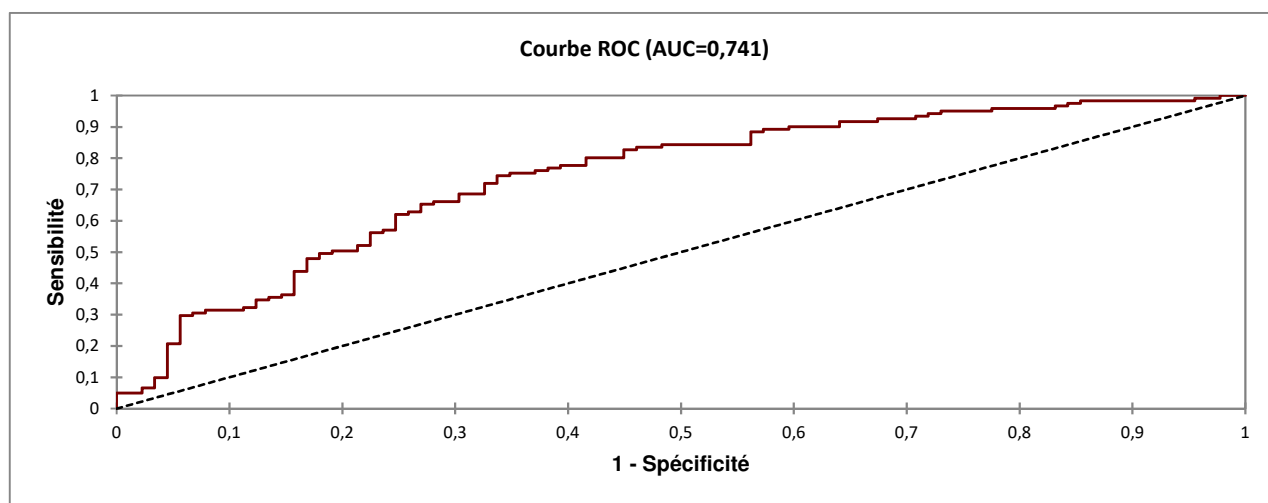
Le logiciel Xlstat fournit un tableau de classification pour l'estimation du modèle. Cette table permet d'appréhender les qualités prévisionnelles du modèle sur l'échantillon en comparant la probabilité estimée pour un individu i d'être $y_i=1$ au seuil arbitraire de 50%.

On voit que parmi les 89 individus pour lesquels $Y_i = 0$, le modèle indique que 46 individus ont une probabilité estimée de retard inférieure à 50%. Dans 51,69% des cas, les non- retards sont donc correctement prévus. On constate également que parmi les 121 individus pour lesquels $Y_i = 1$, le modèle indique que 102 individus ont une probabilité estimée de retard inférieure à 50%. Dans 84,30% des cas, les non-retards sont donc correctement prévus.

Le taux d'erreur de 29,51% ³⁷ est faible. Le pourcentage de correction est donc égal à 70,49%, ce qui signifie que le modèle est vrai dans 70,49% des cas.

La courbe ROC

Graphique 2 - La courbe ROC et l'AUC, aire sous la courbe ROC



Aire sous la courbe: 0,741

La courbe ROC ^{38 39} permet de mettre au point un test qui offre la possibilité de détecter efficacement la survenance d'un événement précis. La sensibilité permet de mesurer à quel point le test est performant. L'aire sous la courbe ou AUC ⁴⁰ est un indice synthétique calculé pour les courbes ROC correspondant à la probabilité qu'un événement positif soit classé comme positif par le test sur l'étendue des valeurs-seuils possibles. On considère habituellement que le modèle est bon dès lors que la valeur de l'AUC est > à 0,7. Dans le cas du présent modèle, la valeur de l'AUC est de 0,741, ce qui autorise à juger la discrimination comme acceptable. La probabilité de distinguer le cas 'présence de retard' du cas 'absence de retard' est de 74,1%.

7. 4 Interprétation économique

Selon Xlstat le modèle s'écrit :

$$\text{Préd(Retard en minutes)} = 1 / (1 + \exp(-(-1,365 + 1,445 * \text{Auto} + 2,856 * \text{Train} + 1,509 * \text{Transport en commun} - 0,625 * \text{Marge de sécurité} / \text{min-2} - 0,995 * \text{Marge de sécurité} / \text{min-3} - 2,307 * \text{Marge de sécurité} / \text{min-4} + 1,123 * \text{bouchon-5} - 1,066 * \text{incidents-1})))$$

L'utilisation de l'auto agit positivement sur la variable Retard, autrement dit les salariés qui utilisent l'auto comme moyen de transport sont susceptibles d'arriver en retard au travail.

Le train est aussi cause manifeste de retard mais le phénomène est bien plus important que celui de l'auto.

³⁷ $(19 + 43) / 210 = 0,2951$

³⁸ <https://www.xlstat.com/fr/solutions/fonctionnalites/courbes-roc> - le 10/6/2018

³⁹ ROC = Receiver Operating Characteristic

⁴⁰ AUC = Area Under the Curve

Bien que le lien de significativité de Transports en commun soit faible, ce moyen de transport se place également dans les risques certains de retard.

La marge de sécurité agit négativement sur la variable retard, donc démarrer avec une avance horaire peut renforcer la garantie d'arriver à l'heure.

Le bouchon est source d'énormément de retard, d'autant plus qu'il est souvent imprévisible.

L'incident 1 agit négativement sur la variable retard. Il s'agit ici des travailleurs qui prétendent ne pas souvent rencontrer d'incidents durant leur trajet. Vu sous cet angle, ils ont plus de chance d'être à l'heure.

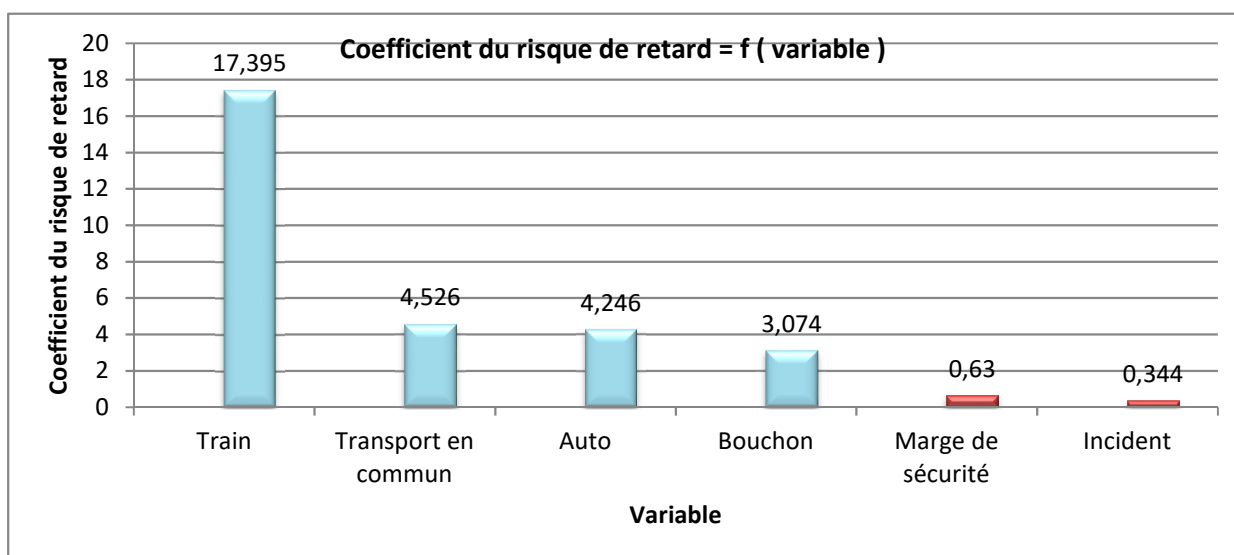
8. Conclusion

"En train ou en voiture, être à l'heure au travail est devenu mission impossible ⁴¹". Une fatalité incontournable ? Les médias commentent largement le phénomène ainsi que les symptômes du malaise. Dans un souci de plus grande objectivité scientifique, le but de cette étude était de quantifier le plus amplement possible la dépendance entre le retard (variable expliquée ou endogène) et les paramètres présumés responsables (variables explicatives ou exogènes), en particulier le moyen de transport, qui induisent ce retard.

Le choix du questionnaire comportait trois inconvénients importants : un nombre important de variables, une nécessité de recourir à deux méthodes (ACM puis Régression Logistique) et le côtoïement de variables de types différents.

D'après le coefficient d'ajustement R^2 (0,260), le présent modèle est de qualité moyenne. Cependant le coefficient test d'ajustement (0,206) de Hosmer-Lemeshow atteste qu'il est bien ajusté. Nous ferons confiance à ce test pour conclure en inspectant les risques de retard.

Graphique n° 3 - Comparaison des risques de retard



Le coefficient de risque de retard est un indicateur précieux. Sur une échelle de valeur, si l'on considère que 1 est l'unité de risque d'arriver en retard, on peut comparer les risques relatifs de chaque cause.

L'auto et le train sont les deux moyens privilégiés de transport. Le risque de retard en train est 4 fois plus élevé que celui de la voiture. L'utilisateur préfère donc la voiture. Mais ce constat mériterait d'être analysé plus finement car l'automobiliste peut aussi préférer utiliser sa voiture pour l'avantage qu'elle lui offre d'arriver confortablement et à proximité immédiate de son lieu de travail (parking), cet avantage lui apparaissant supérieur au risque de retard.

⁴¹ Demeulemeester, Simon. 17/1/2017. Etre à l'heure au travail est devenu mission impossible. Le Vif. 17/1/2017

Les transports en commun (bus, métro) apportent le même niveau de risque que la voiture.

Les bouchons sont souvent imprévisibles, sauf si les travaux sont annoncés par les médias. Si l'utilisateur est prévenu correctement, il peut anticiper : prendre une marge de sécurité plus grande, choisir un autre chemin ou un autre moyen de transport. Confronté à l'imprévu (accidents, intempéries graves), il patiente et estime ce risque à la valeur substantielle de 3,07.

Plus la marge de sécurité est importante, plus le risque de retard diminue. On passe de 0,63 à 0,59. Comme on est < 1 , ce coefficient diminue de toute façon le risque de retard (batonnets bleus). Il en est de même pour le poste incident. Mais dans ce cas, le navetteur semble avoir tiré un enseignement d'expériences antérieures récurrentes. De ce fait, l'incident ne fait plus partie de ses préoccupations ; au contraire il répond positivement face à ce problème.

Si le retard a bien un lien avec les déterminants de la mobilité, d'autres modalités telles que le sexe du travailleur, son âge ou la nature de son occupation professionnelle sont des variables qui n'influencent pas la statistique "retard" et "non retard". Les caractéristiques du travailleur ne sont pas non plus un motif de retard.

Quel enseignement cette étude veut-elle délivrer et à qui ? Au citoyen lambda tout d'abord. Mais peut-on assurer aux gens qu'en lisant cette étude, ils trouveront la solution pour être toujours à l'heure à leur travail ? Non, car donner des critères précis est trop difficile. En revanche, fournir des tendances et des probabilités peut aider les gens à faire les meilleurs choix. Le navetteur disposera de critères pour décider plus adéquatement du choix d'un emploi ou d'un domicile. Si cela n'est pas ou plus réalisable, il est toujours possible d'évaluer la situation existante et de l'améliorer en changeant ses habitudes.

Ensuite remettre un message aux "décideurs". Comment agir ? Il faut déceler en priorité les symptômes du problème, démarche préliminaire logique, pour en chercher des remèdes ⁴². Et enfin clarifier les tendances des embarras causés aux navetteurs pour sensibiliser les décideurs. À ceux-ci de juger de l'urgence de certains investissements et de proposer le choix le plus judicieux pour le confort du citoyen.

⁴² Les maladies semblent bien connues par la presse : le sous-investissement dans les infrastructures, les pannes dues à la défaillance d'un matériel usé ou un manque d'entretien de l'équipement en activité. Mais nous ne participons en rien à ce débat

Limites de l'étude

Si c'était à refaire, quelle serait la section de l'étude qui serait modifiée ? Sans aucun doute, la façon de poser les questions 10 et 12 ⁴³ qui causèrent parfois de l'embarras aux répondants et des difficultés dans le traitement des données.

Reprenons la question 10 (le problème de la question 12 est identique) : *"Pourquoi avoir choisi ce mode de transport principal pour vous rendre au travail ?"*.

Lors de l'élaboration de ces questions, les réponses prévues étaient : 0 = pas d'avis, 1 = pas d'accord, et 5 = d'accord. Les répondants n'ont pas tous réagi dans le sens attendu. Il a donc fallu regrouper certains types de réponses ⁴⁴ (voir détails de l'opération en bas de page).

D'autres ne se sont pas sentis obligés de remplir toutes les cases. Ces vides ont été considérés comme des valeurs manquantes et de ce fait éliminés ⁴⁵ (voir détails de l'opération en bas de page).

Grâce à ces adaptations, l'étude a pu être menée jusqu'à sa finalité telle que programmée.

⁴³ Voir Annexe 5 Détails des questions 10 et 12.

⁴⁴ Malheureusement tous les répondants n'ont pas compris la même chose. De ce fait, des types les réponses différentes ont donc été récoltés : soit, ils répondaient en respectant les trois valeurs permises, soit, ils appliquaient à la lettre une échelle de 5 (0,1. 2. 3. 4,5). Pour résoudre cette ambiguïté et rendre ces données exploitables, sans toutefois biaiser l'information, il a été procédé ainsi : 0=0 pour les "pas d'avis", 1 et 2=1 pour les pas d'accord et enfin 3,4 et 5 =5. Une information orientée dans le même sens a ainsi été conservée.

⁴⁵ Que voulait dire exactement le répondant ? Soit, il voulait dire vide = 0 = pas d'avis ; soit, il était distrait. Comment le savoir ? Il n'était plus possible de redistribuer un questionnaire modifié. Remplacer les vides par des 0 revenait à rajouter artificiellement de l'information. Finalement, la décision a été de considérer ces vides comme des valeurs manquantes et de les traiter comme telles, c'est-à-dire de les supprimer.

Bibliographie

Ouvrages généraux

Bourbonnais, Régis. 2015. *Econométrie - Cours et Exercices corrigés*. Paris : Dunod, 2015.

Esofier, Brigitte et Pagès, Jérôme. 2016 - 5ème édition. *Analyses factorielles simples et multiples*. Paris : Dunod, 2016 - 5ème édition.

Lebart, Ludovic, Piron, Marie et Morineau, Alain. 2006 - 4ème édition. *Statistique exploratoire multidimensionnelle*. Paris : Dunod, 2006 - 4ème édition.

Pagès, Jérôme. 2013. *Analyse factorielle multiple avec R*. Les Ulis : EDP Sciences, 2013.

Articles Internet signés

Baccini, Alain. 2010. Statistique Descriptive Multidimensionnelle. <https://www.math.univ-toulouse.fr/~baccini/zpedago/asdm.pdf>. [En ligne] 2010.

Baudot, Yves. Un exemple d'ACM. <http://www.jybaudot.fr/Analdonnees/exacm.html>. [En ligne]

Bouéroux, Mélanie; Pellé, Baptiste. Impact des transports sur la santé des salariés et des entreprises. <http://www.technologia.fr/blog/wp-content/uploads/2010/03/Enquete-Transports-Technologia.pdf>. [En ligne]

Bouyer, Jean. 2012. La régression logistique en épidémiologie. <https://cel.archives-ouvertes.fr/cel-00124335v2/document>. [En ligne] HAL- archives ouvertes, 2012.

Cornélis Eric. Beldam - L'enquête et ses résultats . https://mobilit.belgium.be/sites/default/files/downloads/enquete_resultats.pdf - Beldam. [En ligne]

Hubert, Jean_Paul. Mobilité quotidienne de Belges. https://books.google.be/books/about/La_Mobilit%C3%A9_quotidienne_des_Belges.html?id=wUH8rUNo3qMC&redir_esc=y. [En ligne]

Husson, François. Analyse des Correspondances Multiples - Agrocampus Ouest. <https://www.youtube.com/watch?v=biHScz3OXbw>. [En ligne]

Institut numérique. Annexe 2 - Résultats de l' ACM. <http://www.institut-numerique.org/annexe-2-resultats-de-lacm-52e75a827c4e6>. [En ligne]

Neji, Sonia et Jigorel, Anne-Hélène. La régression logistique. <https://perso.univ-rennes1.fr/valerie.monbet/ExposesM2/2013/La%20re%CC%81gression%20logistique.pdf>. [En ligne]

Pagès, Jérôme. Analyse des Composantes Multiples - Agrocampus Ouest. <https://www.youtube.com/watch?v=sra70Dz29to>. [En ligne]

Rakotomalala, Ricco. 2017. Pratique de la régression logistique. http://eric.univ-lyon2.fr/~ricco/cours/cours/pratique_regression_logistique.pdf. [En ligne] Université Lumière Lyon 2, 2017.

Toint, Philippe. Mobilité quotidienne en Belgique - MOBEL 1999. <https://mobilit.belgium.be/sites/default/files/downloads/mobel1999.pdf>. [En ligne]

Articles Internet anonymes

<http://pun.be/fr/livre/?GCOI=99993102906770> – La mobilité quotidienne des Belges –le 5-5-2018

<https://mobilit.belgium.be/sites/default/files/downloads/mobel1999.pdf>- Mobilité Quotidienne en Belgique –le 5-5-2018

https://mobilit.belgium.be/sites/default/files/downloads/enquete_resultats.pdf - Beldam -le 5-5-2018

<http://www.levif.be/actualite/belgique/en-train-ou-en-voiture-etre-a-l-heure-au-travail-est-devenu-mission-impossible/article-opinion-600415.html> -le 5-5-2018

<http://spss.espaceweb.usherbrooke.ca/pages/stat-inferentielles/regression-logistique.php>11/6/2018

<https://help.xlstat.com/customer/fr/portal/questions/16253648-odds-ratio> 18/6/2018

<https://www.xlstat.com/fr/solutions/fonctionnalites/courbes-roc> - le 10/6/2018 <https://www.xlstat.com/fr/solutions/fonctionnalites/regression-logistique-pour-reponse-binaires-et-multinomiales-logit-probit>. [En ligne]

<https://www.lecho.be/economie-politique/international/general/bruxelles-une-des-villes-les-plus-embouteillees-d-europe/9864564.html>. [En ligne]

Baccini, Alain. 2010. Statistique Descriptive Multidimensionnelle. <https://www.math.univ-toulouse.fr/~baccini/zpedago/asdm.pdf>. [En ligne] 2010. . [En ligne]

Alice, Kriescher. Le chagrin des Belges. *Telepro* 4/1/2018.

Baccini, Alain. 2010. Statistique Descriptive Multidimensionnelle. <https://www.math.univ-toulouse.fr/~baccini/zpedago/asdm.pdf>. [En ligne] 2010.

Baudot, Yves. Un exemple d'ACM. <http://www.jybaudot.fr/Analdonnees/exacm.html>. [En ligne]

Bouéroux, Mélanie; Pellé, Baptiste. Impact des transports sur la santé des salariés et des entreprises. <http://www.technologia.fr/blog/wp-content/uploads/2010/03/Enquete-Transports-Technologia.pdf>. [En ligne]

Bourbonnais, Régis. 2015. *Econométrie - Cours et Exercices corrigés*. Paris : Dunod, 2015.

Bouyer, Jean. 2012. La régression logistique en épidémiologie. <https://cel.archives-ouvertes.fr/cel-00124335v2/document>. [En ligne] HAL- archives ouvertes, 2012.

Cahiers de l'Observatoire de la mobilité de la Région de Bruxelles-Capitale . https://mobilite-mobiliteit.brussels/sites/default/files/les_pratiques_de_deplacement_a_bruxelles_analyses_appfondies.pdf. [En ligne]

Comprendre la régression logistique. <http://www.em-consulte.com/en/article/842576>. [En ligne] EM-Consulte.

Cornélis Eric. Beldam - L'enquête et ses résultats . https://mobilit.belgium.be/sites/default/files/downloads/enquete_resultats.pdf - Beldam. [En ligne]

Cours de Régression logistique - Université de Lyon. https://help.xlstat.com/customer/fr/portal/articles/2062235-regression-logistique-avec-xlstat?b_id=9283. [En ligne]

Demeulemeester, Simon. 17/1/2017. Etre à l'heure au travail est devenu mission impossible. *Le Vif*. 17/1/2017.

—. **17/1/2017.** Etre à l'heure au travail est devenu mission impossible. <http://www.levif.be/actualite/belgique/en-train-ou-en-voiture-etre-a-l-heure-au-travail-est-devenu-mission-impossible/article-opinion-600415.html>. [En ligne] 17/1/2017.

—. **17/1/2017.** Etre à l'heure au travail est devenu mission impossible. *Le Vif*. 17/1/2017.

Esofier, Brigitte et Pagès, Jérôme. 2016 - 5ème édition. *Analyses factorielles simples et multiples*. Paris : Dunod, 2016 - 5ème édition.

Hubert, Jean_Paul. Mobilité quotidienne de Belges. https://books.google.be/books/about/La_Mobilit%C3%A9_quotidienne_des_Belges.html?id=wUH8rUNo3qMC&redir_esc=y. [En ligne]

Husson, François. Analyse des Correspondances Multiples - Agrocampus Ouest. <https://www.youtube.com/watch?v=bihScz3OXbw>. [En ligne]

Institut numérique. Annexe 2 - Résultats de l' ACM. <http://www.institut-numerique.org/annexe-2-resultats-de-lacm-52e75a827c4e6>. [En ligne]

Jérôme, Pagès. 2013. *Analyse factorielle multiple avec R*. Les Ulis : EDP Sciences, 2013.

Kriesche, Alice. Le chagrin des Belges. *Telepro* 4/1/2018.

Lebart, Ludovic, Piron, Marie et Morineau, Alain. 2006 - 4 ème édition. *Statistique exploratoire multidimensionnelle*. Paris : Dunod, 2006 - 4 ème édition.

Neji, Sonia et Jigorel, Anne-Hélène. La régression logistique. <https://perso.univ-rennes1.fr/valerie.monbet/ExposesM2/2013/La%20re%CC%81gression%20logistique.pdf>. [En ligne]

Pagès, Jérôme. Analyse des Composantes Multiples - Agrocampus Ouest. <https://www.youtube.com/watch?v=sra70Dz29to>. [En ligne]

Pagès, Jérôme. 2013. *Analyse factorielle multiple avec R*. Les Ulis : EDP Sciences, 2013.

Rakotomalala, Ricco. 2017. Pratique de la régression logistique. http://eric.univ-lyon2.fr/~ricco/cours/cours/pratique_regression_logistique.pdf. [En ligne] Université Lumière Lyon 2, 2017.

Régression logistique. <http://spss.espaceweb.usherbrooke.ca/pages/stat-inferentielles/regression-logistique.php>. [En ligne] SPSS à l'UdeS.

Régression logistique. <https://www.xlstat.com/fr/solutions/fonctionnalites/regression-logistique-pour-reponse-binaires-et-multinomiales-logit-probit>. [En ligne] XLSTAT.

Régression Logistique. https://fr.wikipedia.org/wiki/R%C3%A9gression_logistique. [En ligne] Wikipédia.

Toint,Philippe. Mobilité quotidienne en Belgique - MOBEL 1999. <https://mobilit.belgium.be/sites/default/files/downloads/mobel1999.pdf>. [En ligne]

Xlstat - La régression logistique -. <https://www.xlstat.com/fr/solutions/fonctionnalites/regression-logistique-pour-reponse-binaires-et-multinomiales-logit-probit>. [En ligne]

Xlstat - Régression logistique. https://help.xlstat.com/customer/fr/portal/articles/2062235-regression-logistique-avec-xlstat?b_id=9283. [En ligne]

Xlstat. 19/03/2018. Analyse des Correspondances Multiples ACM dans Excel. https://help.xlstat.com/customer/fr/portal/articles/2062224-analyse-des-correspondances-multiples-acm-avec-xlstat?b_id=9283. [En ligne] 19/03/2018.

Articles de presse

Demeulemeester, Simon. 17/1/2017. Etre à l'heure au travail est devenu mission impossible. *Le Vif*

Kriesche, Alice. Dossier mobilité- Le chagrin des Belges. *Telepro* 4/1/2018.

Annexes

Annexe 1 - Tableau des sources de données et des taux de réponse

	Sources	Etudiante n° 1	Etudiante n° 2	Etudiante n°3	Assistant	Total
	Zone prospectée	Namur - Bruxelles	Liège	Arlon	Unamur	
a	# contactés	117	110	90	12	329
b	# refus	44	0	0	3	47
c	# trop tardifs	22	20	5	0	47
d	# réponses retour ⁴⁶	51	90	85	9	235
e	# avec données manquantes					25
f	# exploitables ⁴⁷					210
	La technique de récolte	Envoi par voie électronique Récolte à l'université, aux magasins Carrefour et Trafic, à la Gare, au Centre de Namur.	Envoi par voie électronique Full contact dans son lieu de travail Visite au terrain de football et chez les amis et la famille Contacts téléphoniques.	Tous les questionnaires ont été remplis en présence du collecteur, 20 questionnaires ont été remplis dans un cours spécialisé de droit à l'UCL mais en présence d'un collecteur fiable d'Arlon.		
g	Taux de réponses exploitables ⁴⁸					64%

⁴⁶ d = a - b - c

⁴⁷ f = d - e

⁴⁸ g = f / a

Annexe 2 - Tableau synthétique des variables et de leurs instances dans la Base de Données

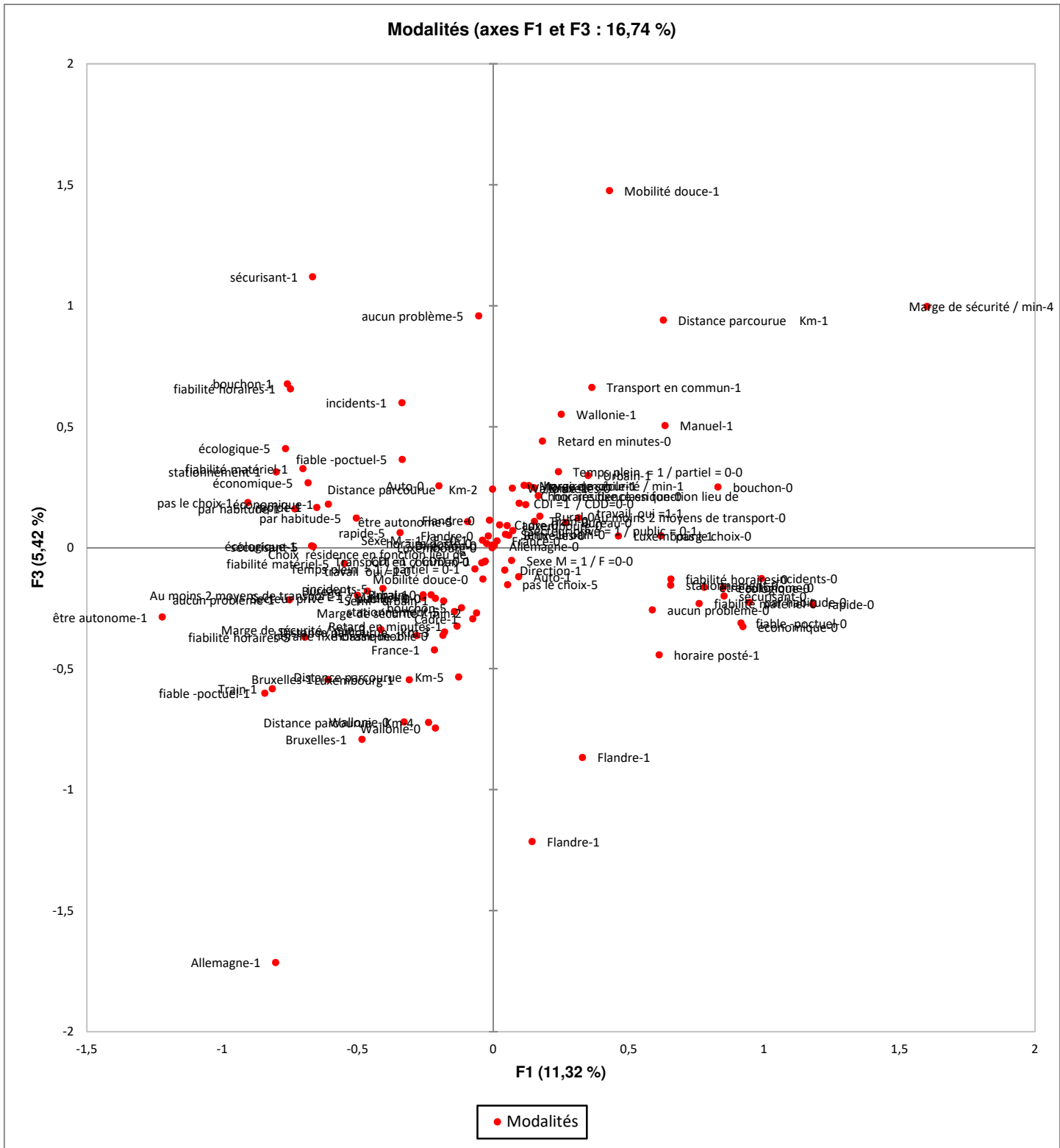
Nom du Groupe sur le tableau brut	N°	Instances possibles pour chaque groupe	N° des instances	Type	Commentaire explicatif
Moyens de transport	1	Auto	1	Qual binaire	L'utilisateur choisit son mode de transport favori en cochant 1. Mais un problème se pose lorsqu'un usager utilise plusieurs moyens de transport (le même jour ou change de moyen selon son humeur !). Dans ce cas on a mis un code (provisoire) de priorité : Moyen principal : code 1 ; moyen secondaire : code 0. 5; moyen tertiaire : code 0. 25
		Bus	2	Qual binaire	
		A pied	3	Qual binaire	
		Métro	4	Qual binaire	
		Train	5	Qual binaire	
		Tram	6	Qual binaire	
		Vélo	7	Qual binaire	
		Autre = Moto	8	Qual binaire	
Retard en minutes / trimestre	2	Retard en minutes /trimestre	9	Quantitative	Il s'agit de la Variable Expliquée. Elle est calculée selon une grille de valeurs discrètes moyennes en minute X la fréquence de ce retard sur un trimestre. Le retard n'est pris en considération qu'après 5 min
Région de travail	3	Bruxelles	10	Qual binaire	L'utilisateur coche 1 dans la zone où il travaille.
		Flandre	11	Qual binaire	
		Wallonie	12	Qual binaire	
		Luxembourg	13	Qual binaire	
Population du lieu de travail	4	Population du lieu de travail	14	Quantitative	Il s'agit du nombre d'habitants de la ville où travaille le citoyen interrogé. Le nombre d'habitants est repris sur la " Liste des communes de Belgique par population" -Wikipédia ⁴⁹
Distance parcourue	5	Distance parcourue	15	Quantitative	La personne interrogée a estimé le nombre de km de son trajet (aller simple) à partir d'une grille de valeurs moyennes discrètes : 2,5/18/40,5/75,5/100 ⁵⁰
Région d'habitation	6	Bruxelles	16	Qual binaire	L'utilisateur coche 1 dans la zone où il habite.
		Flandre	17	Qual binaire	
		Wallonie	18	Qual binaire	
		Luxembourg	19	Qual binaire	
		France	20	Qual binaire	
		Allemagne	21	Qual binaire	
Habitat	7	Rural	22	Qual binaire	L'utilisateur coche 1 pour le type d'habitat où se trouve sa maison.
		Urbain	23	Qual binaire	
		Semi- urbain	24	Qual binaire	
Choix résidence	8	Choix de résidence	25	Qual binaire	L'utilisateur coche 1 dans le cas où il a choisi son lieu de résidence à un endroit favorable pour se rendre rapidement à son travail. S'il change de travail, il n'est pas toujours aisé de déménager.
Raison du mode de transport	9	Pas de choix	26	Qualitative	On demande à l'utilisateur à l'utilisateur de coter chaque qualité selon une échelle de 0 à 5. On veut définir les "goûts " du consommateur.
		Par habitude	27	Qualitative	
		Ecologique	28	Qualitative	
		Economique	29	Qualitative	
		Fiable -ponctuel	30	Qualitative	
		Rapide	31	Qualitative	
		Sécurité	32	Qualitative	
		Autonomie	33	Qualitative	
Marge de sécurité	10	Marge de sécurité au départ	34	Quantitative	Les personnes interrogées ont estimé le nombre de minutes de sécurité prises au départ à partir d'une grille de valeurs moyennes discrètes : 7. 5/20. 5/45. 5/60 ⁵¹ . Beaucoup estiment être suffisamment bien organisées que pour ne pas avoir besoin de cette sécurité.
Type de problème	11	Aucun problème	35	Qualitative	On demande à l'utilisateur à l'utilisateur de coter chaque problème selon une échelle de 0 à 5. On veut repérer les problèmes les plus courants rencontrés par l'utilisateur lors de son déplacement.
		Bouchon	36	Qualitative	
		Stationnement	37	Qualitative	
		Incidents	38	Qualitative	
		Fiabilité matériel	39	Qualitative	
		Fiabilité horaire	40	Qualitative	
		Autre	41	Qualitative	
Type de contrat	12	Type de contrat de travail	42	Qual binaire	Le CDI est noté 1. Le CDD est noté 0
Type d'horaire	13	Posté	43	Qual binaire	L'utilisateur doit coter 1 dans la colonne qui le concerne. On a mis dans "horaire mobile" toutes les particularités comme horaire spécial et horaire sur mesure.
		Classique	44	Qual binaire	
		Mobile	45	Qual binaire	
Temps plein ou partiel	14	Temps plein ou partiel	46	Qual binaire	Le temps plein est coté 1 ; l'horaire partiel est coté 0
Type d'occupation	15	Direction	47	Qual binaire	L'utilisateur doit coter 1 dans la colonne qui le concerne.
		Cadre	48	Qual binaire	
		Bureau	49	Qual binaire	
		Manuel	50	Qual binaire	
Secteur privé /public	16	Secteur privé /public	51	Qual binaire	Si on travaille dans le privé on cote 1, dans le public on cote 0.
Sexe	17	Sexe	52	Qual binaire	M = 0 ; F = 0 /

⁴⁹ https://fr.wikipedia.org/wiki/Liste_des_communes_de_Belgique_par_population

⁵⁰ Ces chiffres sont issus de la moyenne des bornes des plages ; exemple plage entre 31 km et 50 km = (31+50) / 2 = 40,5 km

⁵¹ Ces chiffres sont issus de la moyenne des bornes des plages ; exemple plage entre 11'et 30' = (11+30) / 2 = 20,5'

Annexe 3 - Graphique agrandi des modalités dans ACM



Annexe 4- Tableau des arrivées en retard pour les horaires mobiles, la direction et les cadres

Variables	Nombre total	Nombre d'arrivées en retard	% en retard
Horaire mobile	123	58	47 %
Direction	21	11	52 %
Cadre	51	26	51 %

Dans le nettoyage des données, nous avons éliminé les données manquantes. Dans ce cadre de l'assainissement de la base de données, on pouvait aussi se dire que les variables "Horaire mobile, Direction, Cadre" excluent a priori la possibilité d'être en retard et que ces données sont donc inutiles. Pour cela il aurait fallu que l'on ne trouve aucune arrivée tardive dans ces variables. Or il n'en est rien, comme le montrent les relevés ci-dessus. (cfr annexe 4).

Annexe 5- Détails des questions 10 et 12 extraites du questionnaire de départ

10 Pourquoi avoir choisi ce mode de transport principal pour vous rendre au travail ?									
Parmi les affirmations suivantes , indiquez, sur <u>une échelle de 0 à 5</u> , dans quelle mesure vous êtes d'accord avec celle -ci ?									
pas d'avis		cochez 0							
pas d'accord		cochez 1							
tout à fait d'accord		cochez 5		mettre un chiffre					
Je n'ai pas le choix									
J'ai l'habitude									
Il est écologique.									
Il est économique - le moins coûteux									
Il est fiable et ponctuel									
Il est le plus rapide.									
Il est sécurisant									
Il me permet d'être autonome.									
11 Quelle marge de sécurité (en minutes) prenez-vous avant de vous rendre au travail par rapport à un temps estimé normal pour être certain d'arriver à l'heure?									
				Cochez x					
				entre	5' et 10'				
				entre	11' et 30'				
				entre	31' et 60'				
				entre	plus de 60'				
12 Quel(s) type(s) de problèmes pensez-vous éventuellement rencontrer en vous rendant à votre travail ?									
Parmi les affirmations suivantes , indiquez, sur <u>une échelle de 0 à 5</u> , dans quelle mesure vous êtes d'accord avec celle -ci ?									
pas d'avis		cochez 0							
pas d'accord		cochez 1							
tout à fait d'accord		cochez 5		mettre un chiffre					
Aucun problème									
Bouchons et embouteillage dus à la densité du Trafic									
Difficulté de stationnement									
Incidents exceptionnels (suicide, accident, intempéries importantes, travaux)									
Manque de fiabilité dans le matériel (panne moteur , rupture de caténaire)									
Manque de fiabilité dans les horaires (retard, grève...)									
Autres - mettre un nom									
13 Avez-vous un ?									
				CDD		Cochez x			
				CDI					